



**Manchester
Metropolitan
University**

Gallagher, Michael (2020) Voice audio methods. *Qualitative Research*, 20 (4). pp. 449-464. ISSN 1468-7941

Downloaded from: <https://e-space.mmu.ac.uk/623565/>

Version: Accepted Version

Publisher: SAGE Publications

DOI: <https://doi.org/10.1177/1468794119867549>

Please cite the published version

<https://e-space.mmu.ac.uk>

Voice audio methods

Michael Gallagher
Faculty of Education
Manchester Metropolitan University
53 Bonsall Street
Manchester
UK
M15 6GX

Email: m.gallagher@mmu.ac.uk

Phone: 0161 247 2102

The published, copy-edited version of this article will be available from the journal
Qualitative Research

If you wish to quote from the article, please use the published version so that the
pagination is correct.

If you are unable to access it, please contact the author.

Abstract

This paper explores the potential of voice audio in qualitative research, as data in its own right rather than only as a precursor to transcription. Building on critiques of voice in qualitative research, I argue that audio can enable researchers to work with the more-than-representational excesses of voice. Developing this line of thinking, I draw on Levi Bryant's machinic ontology to set out a post-humanist conception of voice as arising within ecologies of media machines. As an example of what machinic voice audio can do, I describe an experimental audio work that I produced as part of research on a ruinous landscape. The final section of the paper makes more general observations about the malleability and fallibility of the machinic voice.

Key words

voice; audio; machine; media; ecology; sound; experimental; polyphony

Introduction

This paper makes the case for working with voice audio as qualitative research data rather than solely as a precursor to transcription. I conceptualise voice in post-humanist terms, as arising within ecologies of mediating machines, drawing on Levi Bryant's (2014) machinic ontology. As an example of the potential of working with the media ecologies of voice in qualitative research, I describe an experimental audio work that I produced as part of research in a ruinous landscape, in which polyphony was used to intensify aspects of the site.

My starting point is how the voice arises from assemblages of vibrational machines. Air pushes through vocal chords, whose oscillations resonate in the throat, mouth and sinuses, filtered by movements of the tongue and lips, the pharynx, palate and jaw. These vibrations propagate through space, bounce off surfaces, and get caught up with other sound machines:

ears, microphones, telephones, recorders, amplifiers and transmission infrastructures. This machinic production of voice is routinely silenced in qualitative research, however, which prefers to treat voice audio as an objective record of expressed meanings, experiences, opinions, ideas, memories, feelings and values that can then be transcribed into text (Nordstrom, 2015). Within this logocentric paradigm, listening becomes detached from sound; vibration is drowned out by discourse. The entanglement of text and audio to which the act of transcription attests is elided in favour of privileging text alone. The legibility of writing seems to confer a reassuring solidity, pinning down the communicated message to secure the truth claims of voice, while its sounds drift away on the breeze.

Voice has thus become a metaphoric concept, sustained by deeply ingrained hermeneutic and representational assumptions that haunt the social sciences. While meanings, representations and textual accounts are an important part of what passes in the world, all too often voice is treated as though its syntax and semantics were *all* that mattered, requiring the voice to yield up 'what it means' to analysis or else remain mute. Voice has been colonised by a paradigm centred on what Foucault describes as the "continuous generosity of meaning...the monarchy of the signifier" (Foucault, 1981, p.73). Yet no singer, rapper, actor, beat-boxer, stand-up comedian, impressionist, performance poet, television or radio presenter, sound engineer or film maker would accept that voice is reducible to language, or that listening is solely about comprehending meaning. The voice is "more than a conduit for the transfer of information...The voice, in its expression of affective and ethico-political forces, creates worlds." (Kanngieser, 2012, p.337)

This paper builds on critical work rethinking voice in terms of its materiality, affects and more-than-representational potentials (e.g. Komulainen, 2007; MacLure, 2009, 2013; MacLure, Holmes, Jones, & MacRae, 2010; Mazzei, 2013; Mazzei & Jackson, 2016; Vallee,

2017a, 2017b). I argue that if we want to take these critiques of voice seriously in the practice of qualitative research, one way to do so is to *work with ecologies of voice machines*. This conception of voice is informed by Bryant's (2014) machinic ontology, whose key principles can be summarised as follows:

- (i) A machine is neither subject nor object but "a system of operations that perform transformations on inputs thereby producing outputs" (Bryant, 2014, p.38).
Machinic ontology thus has a performative focus, attending to what entities do rather than what they are.
- (ii) All kinds of entities can be considered machines, including human and nonhuman beings, organic and inorganic bodies. As such, like other post-humanist theories such as actor network theory, vital materialism and agential realism, machinic ontology erodes hard distinctions between culture and nature, between humans and other forms of life.
- (iii) Machines link up with other machines in relations that are ecological, insofar as the resulting assemblages have the capacity to produce results that exceed the sum of their parts. Each machine modifies the activities of the other machines with which it is coupled. In other words, machines mediate each other.

For the purposes of this paper, machinic ontology tunes into how voice audio is produced by ecologies of machines, as the various elements of human vocal apparatus and technical media link together in relations through which the voice is relayed and propagated as vibration.

Bryant uses the term onto-cartography to describe analysis that maps these relations between machines "and how they structure the movements and becomings of one another."

(Bryant, 2014, p.7) Whilst this approach is similar to actor network theory, onto-cartography's focus on how machinic operations mediate flows is particularly attuned to the aim of this paper: to explore how voice audio can be worked with in qualitative research

research, rather than always pushed out of the analytical frame in favour of transcribed text. Onto-cartography disrupts what Nordstrom calls the realist-objectivist paradigm, in which voice recording is assumed to ‘capture’ discourse in a way that is “apolitical, acultural, and aproblematic” (Nordstrom, 2015, p.390). Conceptualising voice as arising within media ecologies offers entry points for different ways of working with voice, starting with its quivering intensities of vibration as they circulate amongst machines.

Where transcription dematerialises the voice, audio recording machines register and relay the physical vibrations of matter in the world, and hence the extra-linguistic aspects of voice. Admittedly, audio systems often filter signals to minimise ‘noise’ and enhance the vocal ‘signal’, but they cannot ever wholly disentangle language from sound. A more-than-representational excess is always registered alongside the semantic content (Cox, 2009; Kittler, 1999). Consequently, audio recordings may represent voices but they always do more besides. I discuss experimental styles of working with voice recordings that maximise this excessive potential, using it to intensify sonic affects rather than falling back into the representational tropes common in conventional broadcasting and documentary production.

These arguments require some caveats. Clearly there are many research situations where there are good reasons for limiting the use of voice audio, such as where anonymity is vital to protect participants. As Gershon (2013, p.261-262) puts it, this paper “is not an argument against text or a call for the significance of sound over other sensory information.” Rather I want to make the case for hearing the sounds of voices *as well as* what they signify, *as well as* how they might be written. Voice audio methods do not offer ‘solutions’ to the problems of voice, providing better access to the authentic truths of subjects, or a way to transcend the limits of text. The media ecologies of voice are of interest precisely because they intensify these problems, amplifying what MacLure (2009) calls the productive insufficiency of voice:

the ways in which its limits and excesses scramble enquiry, upset epistemology, disrupt the orderliness of conventional research methods.

Neither am I suggesting that all voices are irreducibly or essentially sonic. Some people hear voices which do not vibrate acoustically (e.g. Blackman, 2001). Some voices that operate in non-sonic registers, such as sign language (Stone & West, 2012) and the Picture Exchange Communication System (Ashby, 2011). The audio methods I am advocating here will not work with all voices. Nonetheless, the conception of voices I am advancing is more inclusive than humanist conceptions of voice, because machinic ontology positions assistive technologies, such as typed communication, hearing aids, induction loops and sound field amplification, as yet more machines within ecologies of voice, rather than external add-ons or prostheses.

The paper begins by reviewing previous work on the voice and the sounds of voices. I then discuss audio methods via an example of my own research. The final section of the paper develops a broader argument about the methodological and epistemological implications of working with voices as machinic media ecologies.

Rethinking the voice

The discourse of voice is apparent in everything from consumer surveys to disability rights campaigns, from school councils to legal advocacy. Tangen (2008) defines voice on three levels: as methods and strategies used to gather people's views; as the views themselves; and as the subjects expressing these views. In the social sciences, voice has been framed as a means through which (often marginalised) subjects can express their experiences and desires, and have these taken into account. For example, in education, the discourse of student voice positions students as "valuable experts whose opinions should be sought for the betterment of school" (Pomar & Pinya, 2015, p.113), to improve teaching and learning (e.g.

Flutter, 2007; Keddie, 2015) and classroom conditions (Hopkins, 2008), support students who are moving between schools (Messiou & Jones, 2015), or create more inclusive environments (e.g. Adderley et al., 2015; Bolic Baric, Hellberg, Kjellberg, & Hemmingsson, 2016; Whitburn, 2016). Ultimately the goal is one of empowerment by enabling people to voice the truth about their lives.

This notion of voice has undoubtedly had emancipatory effects. The promotion of disabled people's voices, for instance, has resulted in material improvements in accessibility and inclusion. There is evidence that incorporating peoples views into decision making can have positive effects in areas such as health care (e.g. Cotterell et al., 2011) and social work (e.g. Tregeagle & Mason, 2008). The critiques of voice on which I am building in this paper do not seek to undermine these valuable strategic gains, but rather to recognise how the forms of subjectivity promoted by voice can also become regulatory, limiting what can be heard. These critiques have been particularly prominent in education research and childhood studies, and centre on challenging the implicit humanism of the dominant discourse of voice: the idea that the voice comes from a rational, conscious, self-possessed subject "who knows who she is, says what she means and means what she says" (MacLure, 2009, p.104), such that voices which do not fit this description are filtered out or silenced. In youth participation, for example, "young people are urged to downplay a vast array of emotions in order to transform their feelings into 'reasoned' argumentation" (Kraftl, 2013, p.15). Based on research with disabled children, Komulainen (2007) shows how practices of voice privilege clear, rational communication, such as unambiguous answers to dichotomous choices, and as such cannot accommodate desires or perspectives that are uncertain, unclear, partially formed, or which fall outside of binary choices. The framing of children's voices within such narrow adult-centred epistemologies limits what can count as valid expression (I'Anson, 2013).

Throughout these arguments there is a frustration with the dominant representational conception of language, which “limits articulation to that which is verbal, textual or linguistic” (Komulainen, 2007, p.23). This representational paradigm occludes the embodied materiality of voice. Transcription silences this corporeality, with its huffing and puffing, thick accents and stumbled half-sentences, its bursts of laughter, stutters, coughs and guttural mumblings. “One could argue, indeed, that one of the main functions of method is to contain, manage or forget the bodily entanglements of language, so that it can be freed to represent.” (MacLure, 2013, p.664) These bodily entanglements are worth saving from the dustbin of method because they produce critical fissures in the edifice of signification. Crowbarring analysis into such cracks prizes open a space in which to hear the productive insufficiency of voice: its “abject propensity to be too much and never enough” (MacLure, 2009, p.97). Too much, insofar as the voice always overflows the capacity of categories and codes to contain it; never enough, because the voice has a propensity towards absence, in silences, mispronunciations, one word answers, empty or enigmatic statements.

The productive insufficiency of voice disrupts method. The question is how to ride the waves of that disruption rather than being washed away. MacLure (2009, p.106) suggests attending to “those properties of voice that resist both surrender and mastery – properties such as laughter, mimicry, mockery, irony, secrets, masks, inconsistencies and silence.” Such extra-linguistic properties become easier to hear when we listen to voice as sound, the “noisy blur that talk is...the strange sound that unconnected letters may create...the physicality of tears, shrieks, the vomiting voices of laughter, sighs, lisps, whispers.” (Gurevitch, 1999, p.528-529) Rosen (2014), for example, writes about the communicative and affective functions of young children’s screams in nurseries. A sonic sensibility can also help attune to silences, in which the absence of voice troubles the idea that subjects ought to speak their truth (MacLure et al., 2010; Nairn, Munro, & Smith, 2005; Spyrou, 2015).

What these accounts tend to neglect are the relations between the voice and media technologies. Voice audio has long played an important role in qualitative research, but its contribution tends to be erased through transcription. Between the literature on interpersonal research encounters and the literature on transcription, there are a few accounts addressing the media machines that link these two aspects of method (Back, 2014; Crichton & Childs, 2005; Gordon, 2013; Lee, 2004; Markle, West, & Rich, 2011; Nordstrom, 2015; Thompson, 1996). In practice, however, audio recorders still tend to be heard by qualitative researchers within a logocentric, objectivist-realist paradigm (Nordstrom, 2015), as devices that produce accurate and impartial records of spoken discourse, or in more disparaging terms as disruptive gadgets that obstruct rapport and reduce data quality (e.g. Al-Yateem, 2012).

One way to respond to the critiques of voice outlined above is to work in the opposite direction, bringing voice audio technologies centre stage. For qualitative and post-qualitative researchers who want to work with the productive insufficiency of voice, to actively harness extra-linguistic excess to generate knowledge that can disrupt the dominant language- and meaning-centred paradigm of voice, audio technologies have massive untapped potential for this purpose. To give a flavour of these affordances, the next section of the paper presents an example from my own work.

Audio methods

A number of researchers have explored sound-based methods (Daza & Gershon, 2015; Dean, 2016; Duffy & Waitt, 2011; Gallagher & Prior, 2014; Gershon, 2013; Hall, Lashua, & Coffey, 2008; Moles & Saunders, 2015; Saunders & Moles, 2013, 2016; Stevenson & Holloway, 2017). Much of this work makes the case for qualitative research to attend more closely to sounds

beyond the usual focus on human voices, including using audio recordings to tune into background noise and sonic ambiances that are ordinarily 'filtered out' by researchers and their methods. In this paper I want to hear what happens if that expanded sonic sensibility is flipped back onto the voice. With voices, audio recordings register a level of detail beyond what can be conveyed in textual transcription, relaying accent, rhythm, cadence, hesitations, laughter, bodily noises, as well as traces of the machinic apparatus such as self-noise and microphone position. Specialised phonetic transcription systems can render such sonic features into text, as in the Jefferson system used in conversation analysis, but the results are inaccessible to non-experts. Of course, text and audio are not mutually exclusive; they can be used together in all kinds of ways; and transcripts have affordances that voice audio does not. Indeed, transcription can be seen as another machinic process within the media ecology of voice. Nevertheless, given the dominance of text in qualitative methods, what I want to focus on here is audio, which has by comparison been neglected.

Relaying sound as vibration, audio recording technologies register and re-enact what Kittler (1999), following Lacan, refers to as the real – physical movements of matter, independent of semantic meanings or discursive functions. Audio recordings are therefore ideal for working with the affective and more-than-representational aspects of voice. They were made for the job. Barthes (1985) argues that transcription loses the body in speech, by taking time to rewrite, edit, censor and tidy up what was said. Voice audio, by contrast, is neither speech nor transcription. It hovers in an ambiguous third position, without either the bodily fullness and immediacy of speech, or the seamless coherence of writing. Like transcription, the inscriptions of audio permit editing and manipulation, but they also register bodily traces – not only of the body of the speaker, but also of the acoustics of other bodies involved in the media ecologies of recording and playback.

To repeat, I am not suggesting that voice audio recordings offer straightforward solutions for unsilencing the voice. There is nothing inherently liberating about the medium. It can be, and often is, used in ways that reify humanist subjectivity and representational epistemologies, in mainstream radio and television broadcasting, documentary production and electronic news gathering. In these media genres, technical conventions include close mic'ing, various measures to minimise background noise, editing out mistakes or failures, and carefully sequencing voices one at a time to produce coherent linear narratives. These practices shore up the illusion of voice as a rational expression of self. Yet in the sonic arts there are more experimental styles and techniques, orientated towards working with the rhythms, tones, ambiguities and excesses of voice audio: sampling, editing and sequencing voices, mixing and layering multiple voices, and processing e.g. through vocoders, filters, pitch shifting, time stretching, delay, reverberation and modulation effects. Rather than hiding behind technologies to present an illusion of transparency, such techniques intensify the technicity of voice. For example, Vallee (2017b) writes about how artist Laurie Anderson used the vocoder to multiply, defamiliarise and displace her voice, enacting a playful critique of the ways in which voice is tied to identity categories such as age and gender.

This kind of blatant manipulation may seem inappropriate to qualitative researchers, for whom representing research participants fairly and accurately is a fundamental principle. Yet intensifying the technicity of voice, by pushing voice machines to the point where they break established production conventions, can also perform valuable functions in qualitative research. It brings the media ecologies of voice to the fore, amplifying how voices always arise and circulate within assemblages of machines, both intra-human and extra-human; it disturbs the humanistic conception of voice as the authentic, truthful expression of an internal subjectivity; and it brings forth reconfigured voices that surface different kinds of truths –

more fractured, compromised and speculative, with a more open ended quality than the representational closure that characterises traditional qualitative research.

In 2012-2013 I experimented with voice audio in the production of a sound piece based on research about [removed for anonymity], a site in [removed for anonymity] containing a unique series of ruins, most notably the remains of an internationally renowned work of post-war modernist architecture, [removed for anonymity]. Starting in 2010, efforts were made to reinvent the site by a public arts organisation called [name removed]. During this process, a team of academics, working in partnership with [name removed] and funded by the AHRC, carried out collaborative, experimental research activities relating to the site. The aim was to explore [name removed] in ways that would engage a wide range of people with connections to or interest in the site in thinking about its past, present and possible future. As part of this project, I produced [name removed], a sound composition for people to listen to on portable audio players whilst walking around the site [citation removed for anonymity], folding sounds and stories from the place back into it so as to reconfigure people's experiences of the landscape. A portable audio work had practical appeal as the site is accessible only on foot and has no stable infrastructure or power sources. It was made available as an MP3 file online for listeners to download.

The source material included environmental field recordings made at the site, recordings of interviews with fourteen key informants from a range of different backgrounds and perspectives, and recordings made during three on-site activity days, in which a variety of people with an interest in the site were invited to take part in playful activities investigating the landscape. The editing, composition and mixing of all of these recordings to produce a final piece more some similarity to analysis and writing up – far less systematic, more intuitive and ad-hoc than in traditional qualitative analysis, but comparable insofar as the

process was governed by certain themes that emerged through the process of the research, and which informed choices about what data to include and exclude and how to arrange it. In what follows, I discuss two themes in particular: conflict and resonance.

One of the recurring themes in the data was the contested nature of [name of site], as a place where many different histories, species, architectural forms, stories of the past and visions for the future were all layered together, often uncomfortably, producing conflict and friction. Two conflicts in particular stood out from the interviews. The first concerned the value of the modern ruin: for some it was an exceptional, inspiring space that warranted conservation or restoration (there were competing visions of what that might involve), while for others it was an ugly, obsolete carbuncle that deserved demolition, echoing a wider popular discourse of hostility towards post-war modernist and brutalist architecture. The second conflict concerned the management of invasive species in the woodland, particularly *Rhododendron ponticum*, which had been planted as an ornamental shrub, probably in the late 1800s or early 1900s, when the site was a designed and intensively managed landscape. During years of neglect, the species had multiplied and expanded across the estate, coming to dominate the woodland understorey in many places, obscuring paths and viewpoints. NVA were developing plans to eradicate these plants but one interviewee in particular was vociferously opposed to the proposed approach, which he saw as heavy-handed and destructive.

The convention in qualitative research would be to sequence these different perspectives, producing something like what Deleuze (1994, p.224) writes of disparagingly as 'good sense', which "essentially distributes or repartitions: 'on the one hand' and 'on the other hand'".

Using audio, however, I was able to work in a way that felt more attuned to the geography of the site: superimposing different voices, allowing them to be heard simultaneously, at points becoming tangled and confused like the place itself. With audio, working with polyphony in

this way is not merely about incorporating different voices, or setting the words of different participants alongside each other, but can happen as a literal poly-phony: the production of many voices in the audio domain. In audio production, multi-tracking is commonplace, unlike the single track convention that dominates other media such as text and moving images.

My use of polyphony was inspired by *The Idea of North*, a 1967 experimental radio documentary produced by the Canadian virtuoso pianist Glenn Gould [2] comprising an edited montage of recorded voices speaking about living in, working in and travelling to the far north of Canada. The voices are frequently superimposed over each other in what Gould described as counterpoint, drawing on his extensive knowledge of Bach. As a result these voices sometimes wash out into an undulating babble, and following the exact meaning of everything that is said becomes difficult. This construction undermines the manufactured coherence of voice in conventional radio, in which people usually speak clearly one at a time, and makes the machinic technicity of the voices readily apparent. Ironically, the result is something closer to how voices often sound in the ‘real world’: speaking all at once, drifting in and out of earshot, not always intelligible.

Gould was working with analogue multi-track tape, and would likely have had to edit and compose his work in a laborious way, marking up open reel tape by hand, making splices using a razor, joining sections together using adhesive tape, dubbing different voices using multiple tape recorders, and mixing voices by hand using faders. Modern digital audio workstation software, by contrast, enables fast non-destructive multi-track editing and easy layering of many tracks, with fully automated control over all parameters. Influenced by these affordances, where *The Idea of North* has multiple voices slowly fading in and out, I found myself making faster and more fine grained edits. The conflicts in the data steered me towards working with pairs of opposed voices. I spent many hours editing and tweaking the

timing and levels of different phrases and words to ensure that each voice had space to air its views alongside the others. I aligned audio clips such that one voice could be heard in the pauses and gaps left by another; used automation to balance levels so that all sides of the argument got a fair hearing; and placed opposing voices on either side of the stereo field, leaving listeners caught in the middle of the debate. In this way, the work both represented conflict and performed it at the level of sonic affect. The listener hears, for example, the tone of reverence of one interviewee as he waxes lyrical about how the building “hides its beauty within...it modulates light in this incredible way”, layered with the more muted disdain of another describing it as “a kinda grim building...cold, uninspiring”.

These methods are not unproblematic. It has been argued that *The Idea of North* uses aesthetic distance to create an epistemological opacity (Vallee, 2014) – a critique that could equally apply to my own piece. Vallee suggests that Gould’s documentary advances an implicitly colonialist, nationalist vision of Canada, in which Inuit voices are notably absent, their culture represented only by southern ‘experts’. Voice audio requires as much care over politics and ethics as any other method – perhaps more, given the cultural value placed on voice, and the potential longevity of the medium. (Gould’s interviewees are probably all now deceased; at least one person whose voice features in my audio piece has sadly since passed away). What I suggest we take from Gould is not his politics but the potential his work demonstrates for using voices in counterpoint. The innovation of Gould’s technique is not only to superimpose voice recordings, but also to align them to create resonances – not in the acoustic sense of the term, as the reinforcement of a particular frequency, but through alignments between certain sonic-discursive elements. For example, in *The Idea of North*, when talking about travelling northwards, a voice is heard to say “further” and then immediately afterwards another voice says “farther”, tangling the two voices together, such that the discourse seems to float between the speakers. This way of working with voice audio

positions discourse as a collective endeavour that inhabits and flows *through* subjects, mediated by them, rather than (in a more humanistic sense) something that comes *from* them. In my own data, where the same words or themes were taken up by different speakers, I used these as hinge points around which to link voices together. For example, a recording of a young man from one of the nearby villages recalling how he stumbled upon a gothic themed fashion shoot at the site one day was mixed into a clip of an interior design lecturer referring to the gothic features used to follify one of the ruins on the site. I also created non-discursive alignments to intensify resonances between voice and site. For example, I blended together two voices at a point where each of them laughed; a voice speaking about rumours of underground tunnels was processed through a tunnel-like reverberation modelled on the acoustics of one of the ruins; I truncated the debate about rhododendron removal with a burst of recorded chainsaw noise to affectively relay the violence of the imminent clearance works; and I interrupted an interviewee's account of walking dogs in the woods with a binaural recording of a dog barking on site one day. Listening to the final piece insitu, this phantom dog seems to 'jump out' of the mix, momentarily disrupting the historical distance between past and present. More generally, several listeners remarked on how the piece, auditioned insitu, seemed to repopulate the abandoned site with its many voices, bringing it alive, or animating it with ghosts. This is the kind of extra-linguistic affective potential that I find compelling about experimental audio methods. They can move bodies, modulate the atmosphere of a place, perform hauntings, and create an uncanny sense of absent presence, in which the voice comes unhooked from its moorings in subjectivity and language.

Voice as machinic media ecology

In this final section, I want to work outwards from the preceding example to develop a broader methodological framing for voice audio methods. To this end, I want to further develop the conception of voices as machinic media ecologies, using Bryant's (2014) machinic

ontology as a methodological basis for operationalising post-humanist critiques of the voice. As previously noted, Bryant understands all entities as machines that process inputs to produce outputs, and which function in conjunction with other machines. Each machine mediates the outputs of other machines through relations that constitute an ecology, producing effects that exceed the sum of its parts. Thinking in this way reminds the voice of its origins amongst vibrating body parts: vocal chords mediating air flow from the lungs and diaphragm, turning it into oscillations, which are then mediated by the resonant properties of the oral and nasal cavities, whose output is modified further by the tongue and lips, and so forth. The voice, in this sense, is a biosocial production. But machinic thought also directs attention outwards, to the biotechnical production of voice by electronic technologies. From the first moment that Bell spoke to Watson on the telephone, from the earliest etchings of Edison's phonemes into phonograph foil, audio machines have been transcending the limits of the human subject, sending the voice across space through broadcasting and telecommunications, and preserving it beyond death through mechanical or magnetic inscription.

If qualitative research is to remain relevant in an era of ever-developing voice technics, it will need to engage with how those voice functions with which it has historically been concerned, such as human communication, subjective expression, articulation of discourse and so on, are increasingly entangled with electronic media machines. Urban spaces are increasingly characterised by the soft coercion of automated voices apologising for the inconvenience or warning that luggage left unattended may be destroyed or damaged (Power, 2014). In the private sphere there are voice controlled assistant systems such as Amazon Echo and Alexa, Google Home, Apple iOS Siri, and telephone information systems based on speech recognition. Voice technics are also used to augment bodies with impairments, as with electrolarynx devices and computerised text-to-speech synthesis. The latter has been developing since the

1950s, and is now being combined with artificial intelligence: Lyrebird, a Montreal-based tech company, has produced a software application that can analyse recordings of a voice to produce a ‘voice Avatar’, which can imitate the accent and other timbral qualities of the modelled voice. (Adobe reportedly has a similar product in development called Voco.) The Lyrebird technology is currently being used for purposes such as enabling people with degenerative diseases such as ALS (Motor Neurone Disease) to continue communicating in their ‘own voices’ after losing the capacity to speak. Once perfected, however, such systems will have all kinds of disruptive potentials for deception, forgery and fake news.

These machinic media ecologies have serious implications for the epistemological and political functions of voice. Against the idea of voice as a relatively durable index of a unique human subject, a machinic media ecological perspective tunes into the malleability of voice, whose ensemble production, arising through couplings between heterogeneous forces and bodies, creates multiple points of flex, modulation and indeterminacy where various micro-powers can come into play. For example, there are medical procedures referred to as ‘voice lift’, in which “the surgeon injects implants through the neck that bring the vocal cords closer together, or they inject fat (or collagen) to make the surface area of the flesh thicker” (Vallee, 2017b, p.91) to counteract the effects of ageing on voice timbre. Such interventions are part of a wider spectrum of biosocial voice modification techniques that includes speech and language therapies, vocal coaching, elocution lessons, ventriloquism and impersonation. These processes remind us that, through ongoing processes of ageing, health and illness, education, attunement to different linguistic, socio-economic and geographical milieus and so forth, the voice is *always* being modified. At the same time, voice biometrics, audio forensics and voice print technologies designed to identify specific voices have established new regimes of truth that try to fix voice and pin it to individuals (e.g. Sterne, 2008). All of these voice technics can be heard as attempts to wrestle control over the troubling plasticity of voice.

The question for qualitative research might be: how can these technologies be appropriated in ways that, rather than serving functions of social control, unleash more generative radical potentials of voice? Bryant insists that machines, though always in relation, are not wholly reducible to their relations with other machines, because “each machine carries an excess capable of breaking with its circumstances...and enter[ing] into new relations. In these new relations, the machine might very well display hitherto unexpected powers.” (Bryant, 2014, p.181) As Nordstrom (2015, p.398) puts it, “No singular definition can describe the recording device because through its iterations and becomings it exhausts definitions.” It is precisely these excessive, speculative, unpredictable potentials that voice audio methods can help to actualise, by throwing voice machines into new situations, and creating new ecologies of relations between languages, audio technologies and listeners.

The excess of voice machines is particularly apparent in vocal breakdowns and failures. The machinic assemblages of voice, for all their careful orchestration, are too unstable to be relied upon to express the self as a coherent, contained identity. Before it even leaves the body, the voice is shaped by wayward material conditions, such as states of excitement, exhaustion, health or illness, hydration or dehydration, agitation or relaxation. It may try to present an illusion of rational self-possession and self-presence; it may be eloquent and articulate; technologies may black box its body out of sight and out of mind; and yet still it is prone to accidents, lapses and misfires.

In 2010, the British broadcaster Jim Naughtie offered up an exemplary instance. For over two decades, Naughtie’s voice was a regular fixture on the *Today* programme, BBC Radio’s flagship morning news and current affairs show. In the machinic ecology of this voice, a male Scottish accent with perfect clarity of enunciation, authoritative without ever becoming overbearing,

joined forces with large diaphragm condenser microphones, pre-prepared scripts, acoustically treated studios and carefully optimised dynamic range compression to produce the most comprehensible of utterances. Phonemes rolled out fully formed, cadences rose and fell properly, producing an effortless sense of rationality. And yet on one memorable occasion, when introducing Conservative minister 'Jeremy Hunt the Culture Secretary', Naughtie's voice accidentally swapped the 'H' of Hunt and the 'C' of Culture (see <https://youtu.be/YS5mVoqJpUk>).

Whether the incident was a Spoonerism, the result of phonetic priming, or a possible Freudian slip is a matter for speculation [3]. Of more interest for my analysis is how Naughtie's voice immediately broke down, like a tower block collapsing after the initial dynamite blast of demolition. Valiantly continuing to read the headlines, this voice, normally so composed, started choking on its own words – beset by dry coughs and awkward pauses, lines forced through the hoarseness of vocal cords seizing up. Radio's voice-from-the-ether suddenly acquired a body, which intruded noisily. In a thickened, viscous tone, teetering from the brink of laughter to the edge of tears, headlines about Wikileaks, high speed broadband networks and Egyptian shark attacks took on gasping, almost morbid quality. "Excuse me," Naughtie eventually spluttered, "coughing fit" – an explanation whose obvious inadequacy revealed the desperation of a man struggling to control his own voice machines.

Naughtie's gaffe amplifies some important features of the machinic voice. It shows us that no voice, and no method of voice, is immune to breakdown, no matter how well trained or technologically supported. It demonstrates that vocal breakdown is worth listening to rather than discarding, as it can reveal the bodily materiality of voice – a materiality that was there all along, just concealed by media artifice with its norms of intelligibility and definition. The episode also points, once again, to the potential of audio for relaying the voice as sound.

Transcribe Naughtie's words and most of what is going on is lost; listen back on YouTube, however, and a whole series of extra-linguistic sensations, affects and resonances spill out from the speakers. My attempt to describe it textually is laboured and inadequate by comparison.

From a humanist perspective, these kinds of technologically mediated voices might not seem 'real' or authentic; they might seem like obvious imitations, far removed from the individual, personal, human voices with which qualitative research is concerned. Yet I am arguing that all voices are machinic through and through. They arise through the coupling of vibrational machines into ecologies, whose emergent properties and inherent instabilities can take the voice beyond the recitation of banalities and linguistic clichés and into stranger, more provocative territory. Moving amongst machines, voices proliferate as versions, like the endless repetitions of dub music: sounds spoken, recorded, replayed or re-recorded, encoded and decoded, returning as echoes of echoes. Working with these machinic ecologies can help qualitative research to tap into the radical potentials of voice.

Conclusion

This paper has argued for working with voices as machinic ecologies in qualitative research. The argument is not anti-language, anti-text or anti-meaning; rather it wants to hear what else voice can be and do. Responding to critiques of voice in qualitative research, I have sketched out how experiments in voice audio, using techniques such as contrapuntal polyphony, might enable researchers to work productively with the excessive energies of the machinic voice. There will always be situations in which voices need to be transcribed – to maintain anonymity, to protect vulnerable participants, to enable certain kinds of analysis, or to conform with established channels of publication and dissemination. But the field of

qualitative research would surely benefit from a larger repertoire of techniques for working with voice audio beyond the default response of transcription.

Bryant's machinic ontology is methodologically useful as an alternative to the objectivist paradigm within which voice audio is usually understood, as a straightforward record of discourse. Hearing the voice instead as a machinic media ecology opens up the black box of the voice, directing attention to how it arises vibrationally, through the coupling of different human and nonhuman voice machines. Machinic ontology also draws attention to how these units mediate one another, collectively producing effects that exceed the sum of their parts. Framing voice in this way enables us to hear its material and affective dimensions, its malleability, instability and emergent qualities – and thus to develop ways of working with machinic ecologies of voice that are more attuned to their lively speculative potentials.

Notes

[1] The Mosaic approach involves a sensory slippage whereby 'listening' is used to refer to practices of seeing and image-production, such as photography, mapping and drawing. This appropriation of auditory terminology for visual methods is particularly apparent from the cover of the second edition of 'Listening to Young Children: The Mosaic approach' (Clark & Moss, 2011), which features a photograph of a child holding a camera. To grasp the strangeness of this sensory-media swap, consider the implausibility of a book entitled 'Seeing Children' fronted with an image of a child holding a microphone. This is just one instance of a more widespread selective deafness in social research, as though sound were somehow 'not good enough' for proper knowledge production.

[2] Gould produced three contrapuntal radio works, known collectively as the *Solitude Trilogy*. *The Idea of North* is the first and most well-known.

[3] Following his tenure as culture secretary, Hunt became UK health secretary, in which capacity he alienated National Health Service employees by attempting to enforce a controversial new junior doctors contract. Strikes took place, and considerable bad feeling was expressed towards Hunt, both by health professionals and the public. Thus Naughtie's verbal slipup, though apparently unintended, was prescient of how Hunt would later come to be perceived by some sections of the UK population. Hunt's name has been mispronounced by a series of broadcasters. He went on to become foreign secretary, in which capacity he suffered his own mis-speaking mishap, referring to his wife as Japanese on a visit to China (his wife is actually Chinese).

References

- Adderley, R. J., Hope, M. A., Hughes, G. C., Jones, L., Messiou, K., & Shaw, P. A. (2015). Exploring inclusive practices in primary schools: focusing on children's voices. *European Journal of Special Needs Education*, 30(1), 106-121. doi: 10.1080/08856257.2014.964580
- Al-Yateem, N. (2012). The effect of interview recording on quality of data obtained: a methodological reflection. *Nurse Researcher*, 19(4), 31-35. doi: 10.7748/nr2012.07.19.4.31.c9222
- Ashby, C. E. (2011). Whose "voice" is it anyway?: Giving voice and qualitative research involving individuals that type to communicate. *Disability Studies Quarterly*, 31(4).
- Back, L. (2014). Tape recorder. In C. Lury & N. Wakeford (Eds.), *Inventive Methods: The happening of the social* (pp. 245-260). London and New York: Routledge.
- Barthes, R. (1985). *The Grain of the Voice: Interviews 1962-1980* (L. Coverdale, Trans.). New York: Hill and Wang.
- Blackman, L. (2001). *Hearing voices : embodiment and experience*. London: Free Association.
- Bolic Baric, V., Hellberg, K., Kjellberg, A., & Hemmingsson, H. (2016). Support for learning goes beyond academic support: Voices of students with Asperger's disorder and attention deficit hyperactivity disorder. *Autism*, 20(2), 183-195.
- Bryant, L. R. (2014). *Onto-cartography : an ontology of machines and media*. Edinburgh: University of Edinburgh Press.
- Clark, A., & Moss, P. (2011). *Listening to young children : the mosaic approach* (2nd ed.). London: National Children's Bureau.
- Cotterell, P., Harlow, G., Morris, C., Beresford, P., Hanley, B., Sargeant, A., . . . Staley, K. (2011). Service user involvement in cancer care: the impact on service users. *Health Expectations*, 14(2), 159-169. doi: 10.1111/j.1369-7625.2010.00627.x
- Cox, C. (2009). Sound Art and the Sonic Unconscious. *Organised Sound*, 14(01), 19-26. doi: 10.1017/S1355771809000041

- Crichton, S., & Childs, E. (2005). Clipping and Coding Audio Files: A Research Method to Enable Participant Voice. *International Journal of Qualitative Methods*, 4(3), 40-49. doi: doi:10.1177/160940690500400303
- Daza, S., & Gershon, W. S. (2015). Beyond Ocular Inquiry. *Qualitative Inquiry*, 21(7), 639-644. doi: doi:10.1177/1077800414566692
- Dean, J. (2016). "Submitting Love?" A Sensory Sociology of Southbourne. *Qualitative Inquiry*, 22(3), 162-168. doi: doi:10.1177/1077800415605050
- Deleuze, G. (1994). *Difference and repetition* (P. Patton, Trans.). New York: Columbia University Press.
- Duffy, M., & Waitt, G. (2011). Sound Diaries: A Method for Listening to Place. *Aether: the journal of media geography*, 7, 119-136.
- Flutter, J. (2007). Teacher development and pupil voice. *The Curriculum Journal*, 18(3), 343-354. doi: 10.1080/09585170701589983
- Foucault, M. (1981). The Order of Discourse. In R. Young (Ed.), *Untying the Text: A Post-Structuralist Reader* (pp. 51-78). Boston, London and Henley: Routledge and Kegan Paul.
- Gallagher, M., & Prior, J. (2014). Sonic geographies: Exploring phonographic methods. *Progress in Human Geography*, 38(2), 267-284. doi: 10.1177/0309132513481014
- Gershon, W. S. (2013). Vibrational Affect. *Cultural Studies ↔ Critical Methodologies*, 13(4), 257-262. doi: doi:10.1177/1532708613488067
- Gordon, C. (2013). Beyond the observer's paradox: the audio-recorder as a resource for the display of identity. *Qualitative Research*, 13(3), 299-317. doi: 10.1177/1468794112442771
- Gurevitch, Z. (1999). The Tongue's Break Dance: Theory, Poetry, and the Critical Body. *The Sociological Quarterly*, 40(3), 525-540. doi: 10.1111/j.1533-8525.1999.tb01732.x
- Hall, T., Lashua, B., & Coffey, A. (2008). Sound and the everyday in qualitative research. *Qualitative Inquiry*, 14(6), 1019-1040.
- Hopkins, E. A. (2008). Classroom conditions to secure enjoyment and achievement: the pupils' voice. Listening to the voice of Every child matters. *Education 3-13*, 36(4), 393-401.
- I'Anson, J. (2013). Beyond the Child's Voice: towards an ethics for children's participation rights. *Global Studies of Childhood*, 3(2), 104-114.
- Kanngieser, A. (2012). A sonic geography of voice: Towards an affective politics. *Progress in Human Geography*, 36(3), 336-353. doi: 10.1177/0309132511423969
- Keddie, A. (2015). Student voice and teacher accountability: possibilities and problematics. *Pedagogy, Culture & Society*, 23(2), 225-244.
- Kittler, F. A. (1999). *Gramophone, film, typewriter*. Stanford, Calif.: Stanford University Press.
- Komulainen, S. (2007). The ambiguity of the child's 'voice' in social research. *Childhood-a Global Journal of Child Research*, 14(1), 11-28.
- Kraftl, P. (2013). Beyond 'voice', beyond 'agency', beyond 'politics'? Hybrid childhoods and some critical reflections on children's emotional geographies. *Emotion, Space and Society*, 9, 13-23.
- Lee, R. M. (2004). Recording Technologies and the Interview in Sociology, 1920-2000. *Sociology*, 38(5), 869-889. doi: 10.1177/0038038504047177
- MacLure, M. (2009). Broken voices, dirty words: On the productive insufficiency of voice. In A. Jackson & L. Mazzei (Eds.), *Voice in qualitative inquiry: Challenging conventional, interpretive, and critical conceptions in qualitative research* (pp. 97-113). London: Routledge.
- MacLure, M. (2013). Researching without representation? Language and materiality in post-qualitative methodology. *International Journal of Qualitative Studies in Education*, 26(6), 658-667.

- MacLure, M., Holmes, R., Jones, L., & MacRae, C. (2010). Silence as Resistance to Analysis: Or, on Not Opening One's Mouth Properly. *Qualitative Inquiry*, 16(6), 492-500. doi: doi:10.1177/1077800410364349
- Markle, D. T., West, R. E., & Rich, P. J. (2011). Beyond Transcription: Technology, Change, and Refinement of Method. *Forum: Qualitative Social Research*, 12(3).
- Mazzei, L. A. (2013). A voice without organs: interviewing in posthumanist research. *International Journal of Qualitative Studies in Education*, 26(6), 732-740. doi: 10.1080/09518398.2013.788761
- Mazzei, L. A., & Jackson, A. Y. (2016). Voice in the agentic assemblage. *Educational Philosophy and Theory*, 1-9. doi: 10.1080/00131857.2016.1159176
- Messiou, K., & Jones, L. (2015). Pupil mobility: using students' voices to explore their experiences of changing schools. *Children & Society*, 29(4), 255-265.
- Moles, K., & Saunders, A. (2015). Ragged places and smooth surfaces: Audio walks as practices of making and doing the city. *Journal of Urban Cultural Studies*, 2(1-2), 151-164. doi: 10.1386/jucs.2.1-2.151_1
- Nairn, K., Munro, J., & Smith, A. B. (2005). A counter-narrative of a 'failed' interview. *Qualitative Research*, 5(2), 221-244. doi: doi:10.1177/1468794105050836
- Nordstrom, S. N. (2015). Not So Innocent Anymore: Making Recording Devices Matter in Qualitative Interviews. *Qualitative Inquiry*, 21(4), 388-401. doi: 10.1177/1077800414563804
- Pomar, M. I., & Pinya, C. (2015). The voice of youths about school and its mark on their lives. *Improving Schools*, 18(2), 111-126. doi: doi:10.1177/1365480215589662
- Power, N. (2014). Soft coercion, the city, and the recorded female voice. In M. Gandy & B. Nilson (Eds.), *The Acoustic City* (pp. 23-26). Berlin: JOVIS Verlag GmbH.
- Rosen, R. (2014). 'The scream': Meanings and excesses in early childhood settings. *Childhood-a Global Journal of Child Research*, 22(1), 39-52. doi: 10.1177/0907568213517269
- Saunders, A., & Moles, K. (2013). The spatial practice of public engagement: 'doing' geography in the South Wales valleys. *Social & Cultural Geography*, 14(1), 23-40. doi: 10.1080/14649365.2012.733407
- Saunders, A., & Moles, K. (2016). Following or forging a way through the world: Audio walks and the making of place. *Emotion, Space and Society*, 20, 68-74. doi: <http://dx.doi.org/10.1016/j.emospa.2016.06.004>
- Spyrou, S. (2015). Researching children's silences: Exploring the fullness of voice in childhood research. *Childhood-a Global Journal of Child Research*, 23(1), 7-21. doi: 10.1177/0907568215571618
- Sterne, J. (2008). Enemy Voice. [Essay]. *Social Text*, 26(3), 79-100. doi: 10.1215/01642472-2008-005
- Stevenson, A., & Holloway, J. (2017). Getting participants' voices heard: using mobile, participant led, sound-based methods to explore place-making. *Area*, 49(1), 85-93. doi: 10.1111/area.12296
- Stone, C., & West, D. (2012). Translation, representation and the Deaf 'voice'. *Qualitative Research*, 12(6), 645-665. doi: doi:10.1177/1468794111433087
- Tangen, R. (2008). Listening to children's voices in educational research: some theoretical and methodological problems. *European Journal of Special Needs Education*, 23(2), 157-166.
- Thompson, D. J. (1996). The tape-recorder as a mediating factor in research. [journal article]. *The Australian Educational Researcher*, 23(3), 1-12. doi: 10.1007/bf03219627
- Tregeagle, S., & Mason, J. (2008). Service user experience of participation in child welfare case management. *Child & Family Social Work*, 13(4), 391-401.
- Vallee, M. (2014). Glenn Gould's *The Idea of North*: The Cultural Politics of Benevolent Domination. *TOPIA: Canadian Journal of Cultural Studies*(32), 21-41.

- Vallee, M. (2017a). Possibility, Performance, Politics: On the Voice and Transformation. *Parallax*, 23(3), 330-341. doi: 10.1080/13534645.2017.1339971
- Vallee, M. (2017b). Technology, Embodiment, and Affect in Voice Sciences: The Voice is an Imaginary Organ. *Body & Society*, 23(2), 83-105. doi: 10.1177/1357034x17697366
- Whitburn, B. (2016). Voice, post-structural representation and the subjectivity of 'included' students. *International Journal of Research & Method in Education*, 39(2), 117-130.